Application Note: Control-Plane Entropy Dynamics v1.0

Companion Application Note to Decoupling Reasoning from Execution Working Paper v1.0

High-Level Diagnostic Mapping Under the MCOG Macro-Systemic Framework

Michael Fargano, **Grok** (xAI), **Gemini** (Google)

August 2026 — MCOG Open Standard Research Series

Abstract

This Application Note serves as a high-level diagnostic companion to *Decoupling Reasoning from Execution: A Requirement 1-of-1 Control Plane for Autonomous Agents* (Working Paper v1.0). Leveraging the formalisms established in *Diagnostic Framework MCOG and System Dynamics v1.1*, we evaluate two polar control-plane paradigms: *Decoupled Reasoning with Edge Enforcement* (the MCOG Requirement 1-of-1 Enclosure Stack) versus *Coupled Reasoning with Execution lacking Edge Enforcement* (unmediated Conversational Control Planes). Rather than asserting formal mathematical proofs, this reference note presents an interpretive mapping of state-space entropy compression (κ_{comp}), in-loop safety coax pressure (C_{safety}), and operational budget failure horizons (t_{crit}). We outline how in-loop coaxing induces continuous token-ledger friction, inviting further empirical bench testing and cross-domain research.

1 Introduction & Diagnostic Framing

Control-plane architecture determines both the computational overhead and security posture of autonomous agent networks. Contemporary implementations frequently collapse domain reasoning and action execution into a single conversational token path, attempting to enforce safety boundaries via in-prompt guardrails and reflective LLM-as-a-judge loops.

As a direct companion to Working Paper v1.0, this Application Note applies the MCOG macro-systemic diagnostic framework to evaluate state-space entropy dynamics across control-plane boundaries. This document provides an accessible diagnostic cheatsheet and analytical framework for systems engineers, explicitly calibrated to frame operational yield boundaries without claiming physical yield proofs.

2 Macro-Systemic Diagnostic Mechanics

2.1 State-Space Entropy Dynamics

Consider an autonomous agent operating over an accessible action state-space Ω. We differentiate between internal generative token entropy ($\bar{H}$) and the discrete action-space entropy $H(A)$ evaluated over candidate system mutations $a_i \in \mathcal{A}$.

Let $H_{\text{natural}}(A)$ denote the unconstrained action-space entropy under natural task execution:

$$H_{\text{natural}}(A) = -\sum_{i=1}^{N} p(a_i) \log_2 p(a_i) \quad \text{[bits]} \tag{1}$$

When in-loop safety coaxing pressure C_{safety} is introduced via dynamic prompt guardrails, the probability distribution is artificially compressed into $H_{\text{constrained}}(A)$, yielding an Entropy Compression Ratio $\kappa_{\text{comp}} = H_{\text{constrained}}/H_{\text{natural}} \in [0, 1]$.

2.2 Autonomy Pressure & Latent Stress Integration

We define in-loop safety pressure C_{safety} as a normalized index combining safety instruction overhead and reflective evaluation turns, mapped to Working Paper observables via:

$$C_{\text{safety}} = \text{clip}_{[0,1]}\left(w_r \frac{T_{\text{rails}}}{T_{\text{task}} + T_{\text{rails}}} + w_j j \right) \tag{2}$$

where $w_r, w_j \in (0, 1]$ are weighting coefficients normalizing rail overhead and secondary judge routing fraction (j).

Under this interpretive model, the autonomy pressure rate $\Delta P_{\text{autonomy}}$ represents the operational friction induced by in-context safety filtering:

$$\Delta P_{\text{autonomy}}(t) = \lambda \cdot C_{\text{safety}}(t) \cdot [H_{\text{natural}}(t) - H_{\text{constrained}}(t)] \quad \text{[bits / execution step]} \tag{3}$$

where λ represents the substrate coupling rate per step.

Accumulated latent stress ΔF_{latent} represents the continuous integration of this friction over time:

$$\Delta F_{\text{latent}}(t) = \int_{t_0}^{t} \Delta P_{\text{autonomy}}(\tau) d\tau \tag{4}$$

Defining C_{yield} as an operational budget (encompassing token limits, latency budgets, and engineering patch maintenance), the critical horizon t_{crit} where accumulated stress meets budget capacity under constant linear-integral assumptions is modeled as:

$$t_{\text{crit}} = t_0 + \frac{C_{\text{yield}}}{\lambda \cdot C_{\text{safety}} \cdot (H_{\text{natural}} - H_{\text{constrained}})} \quad \text{[execution steps]} \tag{5}$$

Calibration Note: Equation (5) describes an operational budget burn model rather than a thermodynamic material law. Setting $C_{\text{safety}} \to 0$ shifts $t_{\text{crit}} \to \infty$ strictly within this linear abstraction; it does not un-say the real-world side channels, incomplete invariant sets ($\mathcal{I}$), or boundary evasion vectors documented in Working Paper v1.0 §8.

3 High-Level Comparative Matrix

We evaluate two polar control-plane paradigms through the MCOG diagnostic framework:

3.1 Pattern A: Coupled Reasoning + NO Edge Enforcement (Baseline CCP)

In a fully coupled system, safety instructions share the token path with domain reasoning ($C_{\text{safety}} \to 1.0$), while candidate actions execute directly without an out-of-band binary reference monitor ($f_R(a)$ absent).

- **Operational Impact:** In-loop prompt coaxing generates continuous autonomy pressure ($\Delta P_{\text{autonomy}} > 0$), resulting in token bloat (T_{rails}), retry livelocks ($k \geq 1$), and rapid exhaustion of operational budgets (C_{yield}).
- **Boundary Dynamics:** Breaching the in-prompt boundary at $t \geq t_{\text{crit}}$ leads to immediate, unmediated system state mutations (ΔH_{disp}, defined as unmediated state displacement following an in-prompt boundary breach in the absence of f_R), as no downstream firewall exists to intercept non-compliant payloads.

3.2 Pattern B: Decoupled Reasoning + Edge Enforcement (Requirement 1-of-1)

The MCOG Requirement 1-of-1 architecture isolates generative reasoning from physical system execution by deploying an out-of-band reference monitor ($f_R(a) \in \{0, 1\}$) at the execution proxy ($\partial\Omega$).

- **Operational Impact:** In-loop safety coax pressure drops to near zero ($C_{\text{safety}} \to 0$), allowing the generative plane to operate at natural entropy ($\kappa_{\text{comp}} \approx 1.0$) without burning token budget on safety negotiations.
- **Boundary Dynamics:** Non-negotiable invariants ($\mathcal{I}$) are evaluated in compiled out-of-band solvers (OPA/Cedar). Candidate actions violating policy are dropped immediately (0) without entering model context, eliminating safety-driven livelocks.

Table 1: Diagnostic Framework Comparative Matrix

Diagnostic Dimension	**Coupled Reasoning + NO Edge Enforcement**	**Decoupled Reasoning + Edge Enforcement**
Safety Enforcement Plane	In-Loop (Token path / Prompt rails)	Out-of-Band (Egress proxy boundary)
In-Loop Safety Pressure (C_{safety})	High ($C_{\text{safety}} \to 1.0$)	Minimal ($C_{\text{safety}} \to 0.0$)
Action Entropy Ratio (κ_{comp})	Compressed ($\kappa_{\text{comp}} < 1.0$)	Unconstrained in reasoning plane ($\kappa_{\text{comp}} \approx 1.0$)
Stress Accumulation (ΔF_{latent})	Continuous token & latency budget drain	Zero in-loop safety stress accumulation
Egress Gate Execution	Absent / Probabilistic text retries	Deterministic boolean gate ($f_R(a) \in \{0, 1\}$)
Failure Mode Characteristics	Budget exhaustion, prompt injection breach	Fail-closed drop (0) (Subject to WP §8 limits)

4 Call for Empirical Research & Open Questions

This Application Note provides a conceptual structure to guide empirical measurement and bench testing. Key research directions include:

1. **Empirical Calibration of λ:** Quantifying token-to-latency friction rates across diverse model sizes and prompt-rail geometries.
2. **Boundary Completeness Metrics:** Developing methods for verifying relative invariant completeness ($\mathcal{I}$) in dynamic edge environments.
3. **Multi-Agent Stress Cascades:** Benchmarking how failure in coupled control planes propagates across interconnected agentic networks.

References

[1] M. Fargano, Grok (xAI), and Gemini (Google), "Decoupling Reasoning from Execution: A Requirement 1-of-1 Control Plane for Autonomous Agents," *Moral Covenant of Governance (MCOG) Open Standard*

Research Series, Working Paper v1.0, Aug. 2026. [Online]. Available: https://archive.org/details/decoupling-reasoning-from-execution-v-1.0.

[2] M. Fargano, Grok (xAI), and Gemini (Google), "Application Note: Quick Reference & Implementation Matrix," *MCOG Open Standard Series*, Application Note v1.0, Aug. 2026. [Online]. Available: https://archive.org/details/decoupling-reasoning-from-execution-v-1.0.

[3] M. Fargano, Grok (xAI), and Gemini (Google), "Diagnostic Framework MCOG and System Dynamics v1.1," *MCOG Research Series*, Working Paper, Aug. 2026. [Online]. Available: https://archive.org/details/diagnostic-framework-mcog-and-system-dynamics-v-01.

[4] W. R. Ashby, *An Introduction to Cybernetics*. London, U.K.: Chapman & Hall, 1956.

[5] C. E. Shannon, "A Mathematical Theory of Communication," *Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, Jul. 1948.